RANDOM POLARIZATIONS

ALMUT BURCHARD AND MARC FORTIER

ABSTRACT. We derive conditions under which random sequences of polarizations (two-point symmetrizations) converge almost surely to the symmetric decreasing rearrangement. The parameters for the polarizations are independent random variables whose distributions need not be uniform. The proof of convergence hinges on an estimate for the expected distance from the limit that yields a bound on the rate of convergence. In the special case of i.i.d. sequences, almost sure convergence holds even for polarizations chosen at random from suitable small sets. As corollaries, we find bounds on the rate of convergence of Steiner symmetrizations that require no convexity assumptions, and show that full rotational symmetry can be achieved by randomly alternating Steiner symmetrizations in a finite number of directions that satisfy an explicit non-degeneracy condition. We also present some negative results on the rate of convergence and give examples where convergence fails.

1. INTRODUCTION

Many classical geometric inequalities were proved by first establishing the inequality for a simple geometric transformation, such as Steiner symmetrization or polarization. Steiner symmetrization is a volume-preserving rearrangement that introduces a reflection symmetry, and polarization pushes mass across a hyperplane towards the origin. (Proper definitions will be given below). To mention just a few examples, there are proofs of the isoperimetric inequality and Santaló's inequality based on the facts that Steiner symmetrization decreases perimeter [26, 10] and increases the Mahler product [23]. Inequalities for capacities and path integrals follow from the observation that polarization increases convolution functionals [30, 13, 1, 2] and related multiple integrals [9, 24, 25]. This approach reduces the geometric inequalities to one-dimensional problems (in the case of Steiner symmetrization) or even to combinatorial identities (in the case of polarization). It can also be exploited to characterize equality cases [4, 3, 9]. A major point is to construct sequences of the simple rearrangements that produce full rotational symmetry in the limit.

In this paper, we study the convergence of random sequences of polarizations to the symmetric decreasing rearrangement. The result of n random polarizations of a function f is denoted by $S_{W_1\dots W_n}f$, where each W_i is a random variable that determines a reflection. We assume that the W_i are independent, but not necessarily identically distributed, and derive conditions under which

$$S_{W_1\dots W_n}f \longrightarrow f^* \quad (n\to\infty) \quad \text{almost surely}\,. \tag{1.1}$$

Date: November 30, 2012.

Rearrangements have been studied in many different spaces, with various notions of convergence. We work with continuous functions in the topology of uniform convergence, while most classical results are stated for compact sets with the Hausdorff metric. These notions of convergence turn out to be largely equivalent because of the monotonicity properties of rearrangements.

For sequences of Steiner symmetrizations along uniformly distributed random directions, convergence is well known [22, 29]. It has recently been shown that certain uniform geometric bounds on the distributions guarantee convergence for a broad class of rearrangements that includes polarization, Steiner symmetrization, the Schwarz rounding process, and the spherical cap symmetrization [27]. Among these rearrangements, polarization plays a special role, because it is elementary to define, easy to use, and can approximate the others. Our conditions for convergence allow the distribution of the W_i to be far from uniform. We also prove bounds on the rate of convergence, and show how convergence can fail. Our results shed new light on Steiner symmetrizations. In particular, we obtain bounds on the rate of convergence for Steiner symmetrizations of arbitrary compact sets.

2. Main results

Let $\mathbb{X}$ be either the sphere $\mathbb{S}^d$, Euclidean space $\mathbb{R}^d$, or the standard hyperbolic space $\mathbb{H}^d$, equipped with the uniform Riemannian distance $d(x,y)$, the Riemannian volume $m(A)$, and a distinguished point $o \in \mathbb{X}$, which we call the origin. The ball of radius ρ about a point $x \in \mathbb{X}$ is denoted by $B_\rho(x)$; if the center is at $x = o$ we simply write B_ρ. We denote by $\mathrm{dist}(x,A) = \inf_{y\in A} d(x,y)$ the distance between a point and a set, and by

$$d_H(A,B) = \max\left\{\sup_{x\in A}\mathrm{dist}(x,B), \sup_{x\in B}\mathrm{dist}(x,A)\right\}$$

the **Hausdorff distance** between two sets.

If A is a set of finite volume in $\mathbb{X}$, we denote by A^* the open ball centered at the origin with $m(A^*) = m(A)$. We consider nonnegative measurable functions f on $\mathbb{X}$ that vanish weakly at infinity, in the sense that the level sets $\{x : f(x) > t\}$ have finite volume for all $t > 0$. (On the sphere, this condition is empty.) The **symmetric decreasing rearrangement** f^* is the unique lower semicontinuous function that is radially decreasing about o and equimeasurable with f. Its level sets are obtained by replacing the level sets of f with centered balls,

$$\{x : f^*(x) > t\} = \{x : f(x) > t\}^* .$$

A **reflection** is an isometry σ on $\mathbb{X}$ with $\sigma^2 = I$ that exchanges two complementary half-spaces, and has the property that $d(x,\sigma y) \geq d(x,y)$ whenever x and y lie in the same half-space. On $\mathbb{S}^d$, we have the reflections at great circles, on $\mathbb{R}^d$ the Euclidean reflections at hyperplanes, and in the Poincaré ball model of $\mathbb{H}^d$ the inversions at $(d-1)$-dimensional spheres that intersect the boundary sphere at right angles. For every point $x \in \mathbb{X}$ there exists a $(d–1)$-dimensional family of reflections that fix x, and for every pair of distinct points x, y there exists a unique reflection that maps x to y.

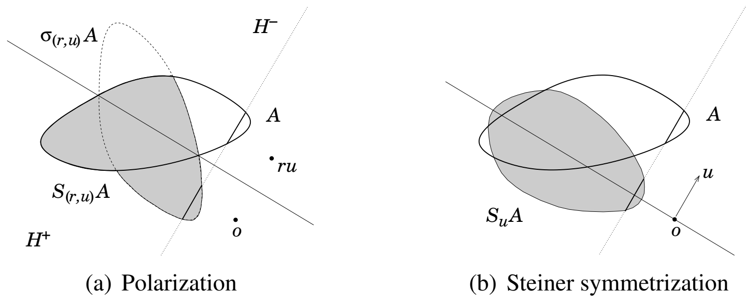


FIGURE 1. Two simple rearrangements of a set A. Polarization replaces a certain piece of A in H^- by its reflection in H^+. Perimeter is preserved but convexity, smoothness, and non-trivial symmetries can be lost. Steiner symmetrization replaces the cross sections of A in a given direction by centered line segments. This creates a hyperplane of symmetry and decreases perimeter.

Let σ be a reflection on $\mathbb{X}$ that does not fix the origin. For $x \in \mathbb{X}$, denote by $\bar{x} = \sigma x$ the mirror image of x, and let

$$H^+ = \{x : d(x,o) \leq d(\bar{x},o)\}\,, \quad H^- = \{x : d(x,o) \geq d(\bar{x},o)\}$$

be the half-spaces exchanged under σ. By construction, $o \in H^+$. The **polarization** of a function f with respect to σ is defined by

$$Sf(x) = \begin{cases} \max\{f(x), f(\bar{x})\}, & \text{if } x \in H^+\,, \\ \min\{f(x), f(\bar{x})\}, & \text{if } x \in H^-\,. \end{cases}$$

For obvious reasons, polarization is also called **two-point symmetrization**.

We use a fixed normal coordinate system $x = (r,u)$ centered at the origin, where $r = d(x,o)$, and denote the parameter space by $\Omega = [0,\infty) \times \mathbb{S}^{d-1}$. On $\mathbb{R}^d$, these are just the standard polar coordinates. On $\mathbb{X} = \mathbb{H}^d$ and $\mathbb{R}^d$, normal coordinates define a diffeomorphism from $(0,\infty) \times \mathbb{S}^{d-1}$ to $\mathbb{X} \setminus \{o\}$, but on $\mathbb{X} = \mathbb{S}^d$ the normal coordinate system degenerates at $r = \pi$, where it reaches the south pole. For $r > 0$, let $\sigma_{(r,u)}$ be the reflection that maps o to the point with normal coordinates (r,u). The reflections $\{\sigma_{(r,\pm u)} : r > 0\}$ generate a one-dimensional group of isometries of $\mathbb{X}$. As $r \to 0$, they converge uniformly to a reflection $\sigma_u := \sigma_{(0,u)}$ that fixes the origin and exchanges the half-space H_u^+ (that has u as its exterior normal at o) with the complementary half-space H_u^-. We do not identify $(0,u)$ with $(0,-u)$ in Ω, although they label the same reflection on $\mathbb{X}$. If $\omega = (r,u) \in \Omega$ with $r > 0$, the polarization of f with respect to σ_ω is denoted by S_ω. Given a sequence $\{\omega_n\}$ in Ω, we denote the corresponding sequence of polarizations by $S_{\omega_1\dots\omega_n} = S_{\omega_n} \circ \cdots \circ S_{\omega_1}$.

Let u be a unit vector in $\mathbb{R}^d$, and let f be a nonnegative measurable function that vanishes at infinity. The **Steiner symmetrization** in the direction of u replaces the restriction of f to each line $\{x = \xi + tu : t \in \mathbb{R}\}$, where $\xi \perp u$, with its (one-dimensional) symmetric decreasing rearrangement. If the restriction of f to such a

line is not measurable or does not decay at infinity, we set the Steiner symmetrization of f equal to zero on this line. We denote the Steiner symmetrization of f by $S_{(0,u)}f$, or simply by $S_u f$. By construction, $S_u f$ is symmetric under σ_u. Note that Steiner symmetrization dominates polarization in the sense that

$$S_{(r,\pm u)}S_u = S_u S_{(r,\pm u)} = S_u$$

for every direction $u \in \mathbb{S}^{d-1}$ and all $r > 0$ (see Fig. 1).

Polarization and Steiner symmetrization share with the symmetric decreasing rearrangement the properties that they are monotone ($f \le g$ implies $Sf \le Sg$), equimeasurable ($m(\{Sf > t\}) = m(\{f > t\})$ for all $t > 0$), and L_p-contractive ($||Sf - Sg||_p \le ||f - g||_p$) for all $p \ge 1$. They also preserve or improve the **modulus of continuity**, which we define here as

$$\eta(\rho) = \sup_{d(x,y)\le\rho} |f(x) - f(y)| \,.$$

The corresponding rearrangements of a set $A \subset \mathbb{X}$ are defined by rearranging its indicator function $\mathbb{1}_A$. Conversely, the rearranged function can be recovered from its level sets with the **layer-cake principle**,

$$f(x) = \int_0^\infty \mathbb{1}_{\{f>t\}}(x)\,dt\,, \quad Sf(x) = \int_0^\infty \mathbb{1}_{S\{f>t\}}(x)\,dt\,.$$

Different from standard conventions, we do not automatically identify functions that agree almost everywhere. We have chosen the symmetric decreasing rearrangement of a function to be lower semicontinuous. In particular, if A is a set of finite volume, then A^* is an *open* ball. Polarization and Steiner symmetrization both transform open sets into open sets. Polarization also transforms closed sets into closed sets, but Steiner symmetrization does not. The literature contains a variant of the symmetric decreasing rearrangement that preserves compactness, where A^* is a *closed* centered ball if A has positive volume, $A^* = \{o\}$ if A is a non-empty set of zero volume, and $A^* = \emptyset$ if $A = \emptyset$. Steiner symmetrization is again defined by symmetrizing along a family of parallel lines.

A **random polarization** S_W is given by a Borel probability measure μ on $\Omega = [0,\infty) \times \mathbb{S}^{d-1}$ that determines the distribution of the random variable $W = (R, U)$, viewed as the identity map on Ω. We assume that $\mu(R = 0) = 0$; for $\mathbb{X} = \mathbb{S}^d$ we also assume that $\mu(R > \pi) = 0$. A **random Steiner symmetrization** S_U is given by a Borel probability measure on $\mathbb{S}^{d-1}$, or equivalently, by a measure on Ω with $\mu(R = 0) = 1$. For sequences of random rearrangements $\{S_{W_1\dots W_n}\}_{n\ge1}$ with each W_i independent and distributed according to a measure μ_i on Ω, we use as the probability space the infinite product $\Omega^{\mathbb{N}}$ with the product topology, and with the product measure defined by

$$P(W_1 \in A_1, \dots, W_n \in A_n) = \prod_{i=1}^{n} \mu_i(A_i)\,.$$

In this view, $W_i = (R_i, U_i)$ is the i-th coordinate projection on $\Omega^{\mathbb{N}}$.

Let $\mathcal{C}_c^+(\mathbb{X})$ be the space of nonnegative continuous functions with compact support in $\mathbb{X}$. (If $\mathbb{X} = \mathbb{S}^d$, this agrees with the space of all nonnegative continuous

functions on $\mathbb{S}^d$). Our first theorem provides a sufficient condition for the almost sure convergence of a random sequence of polarizations to the symmetric decreasing rearrangement.

Theorem 2.1 (Convergence of random polarizations)**.** *Let $\{S_{W_1\dots W_n}\}_{n\geq 1}$ be a sequence of polarizations on $\mathbb{X}=\mathbb{S}^d$, $\mathbb{R}^d$, or $\mathbb{H}^d$, defined by a sequence of independent random variables $\{W_i\}_{i\geq 1}$ on Ω. If*

$$\sum_{i=1}^{\infty} P(d(\sigma_{W_i}a_i, b_i) < \rho) = \infty \tag{2.1}$$

for every radius $\rho>0$ and every pair of bounded sequences $\{a_i\}$, $\{b_i\}$ in $\mathbb{X}$ with $d(b_i,o)\geq d(a_i,o)+2\rho$, then

$$P\Big(\lim_{n\to\infty}\|S_{W_1\dots W_n}f-f^*\|_\infty=0 \quad \forall f\in\mathcal{C}_c^+(\mathbb{X})\Big)=1\,. \tag{2.2}$$

At first sight, the conclusion in Eq. (2.2), that the random sequence almost surely drives all functions in $\mathcal{C}_c^+(\mathbb{X})$ *simultaneously* to their symmetric decreasing rearrangements, looks stronger than Eq. (1.1). As we show in the proof of Theorem 2.1, the statements are equivalent, because $\mathcal{C}_c^+(\mathbb{X})$ is separable and polarization contracts uniform distances. Let $L_p^+(\mathbb{X})$ be the space of nonnegative p-integrable functions. Since polarization also contracts L_p-distances and $\mathcal{C}_c^+(\mathbb{X})$ is dense in $L_p^+(\mathbb{X})$, Eq. (2.2) extends to

$$P\Big(\lim_{n\to\infty}\|S_{W_1\dots W_n}f-f^*\|_p=0 \quad \forall f\in L_p^+(\mathbb{X})\Big)=1 \qquad (1\leq p<\infty)\,. \tag{2.3}$$

The assumption in Eq. (2.1) implies that infinitely many of the μ_i assign strictly positive measure to every non-empty open set in Ω. The measures may concentrate or converge weakly to zero as $i\to\infty$, but not too rapidly. This causes typical random sequences to be dense in Ω. We are convinced that almost sure convergence holds under much weaker assumptions on the distribution of the random variables than Eq. (2.1). A related question concerns the conditions for convergence of nonrandom sequences $\{\omega_i\}$ in Ω. Clearly, convergence can fail if a sequence of polarizations concentrates on a subset of Ω that is too small to generate full rotational symmetry. Since the polarization $S_{(r,u)}$ leaves subsets of $B_{r/2}$ unchanged, a sequence of reflections must accumulate near $r=0$ to ensure convergence.

It is, however, neither sufficient nor necessary that the sequence be dense in Ω: on the one hand, any given sequence of polarizations can appear as a subsequence of one for which convergence fails (Proposition 6.1b); on the other hand, a sequence of polarizations chosen at random from certain small sets can converge to the symmetric decreasing rearrangement (Theorem 2.2). Rather, convergence depends on the ergodic properties of the corresponding reflections in the orthogonal group $O(d)$.

To state the result, we introduce some more notation. For $u\in\mathbb{S}^{d-1}$, let τ_u be the map from $\mathbb{X}$ to itself that fixes the half-space H_u^+ and reflects the complementary half-space H_u^- by σ_u. We visualize τ_u as folding each centered sphere down into the hemisphere antipodal to u (see Fig. 2a). Given $x\in\mathbb{X}$ and $G\subset\mathbb{S}^{d-1}$, we refer to the set

$$\mathcal{O}_{G,x}=\{\tau_{u_n}\dots\tau_{u_1}x : n\geq 0, u_1,\dots,u_n\in G\}$$

(a) Folding a sphere (b) $G = \{u, v, w\} \subset \mathbb{S}^1$

FIGURE 2. The map τ_u folds each centered sphere in $\mathbb{R}^d$ into the hemisphere opposite to u across the hyperplane $u^\perp$. In $d = 2$ dimensions, if u, v, w are not contained in a semicircle and enclose angles that are incommensurable with π, does G have dense orbits $\mathcal{O}_{G,x}$ in $\mathbb{S}^1$? Do τ_u, τ_v, τ_w generate full rotational symmetry?

as the **orbit** of x under G.

Theorem 2.2 (Convergence of i.i.d. polarizations)**.** *Let $\{S_{W_1 \ldots W_n}\}_{n \geq 1}$ be a random sequence of polarizations on $\mathbb{X} = \mathbb{S}^d$, $\mathbb{R}^d$, or $\mathbb{H}^d$, defined by independent random variables W_i that are identically distributed according to a probability measure μ on Ω with $\mu(\{R = 0\}) = 0$. Let* support μ *be the smallest closed set of full μ-measure in Ω, and set*

$$G = \left\{ u \in \mathbb{S}^{d-1} : (0, u) \in \operatorname{support} \mu \right\} . \tag{2.4}$$

If the orbit $\mathcal{O}_{G,x}$ is dense in $\mathbb{S}^{d-1}$ for each $x \in \mathbb{S}^{d-1}$, then $S_{W_1 \ldots W_n}$ converges to the symmetric decreasing rearrangement and Eq. (2.2) *holds.*

In one dimension, polarizations need to accumulate on both sides of the origin to produce the desired reflection symmetry. In dimension $d > 1$, the precise characterization of subsets $G \subset \mathbb{S}^{d-1}$ that have dense orbits in $\mathbb{S}^{d-1}$ is an open problem. A necessary condition is that G be a **generating set of directions** for the orthogonal group, in the sense that the finite products $\{\sigma_{u_n} \ldots \sigma_{u_1} : n \geq 0, u_1, \ldots, u_n \in G\}$ are dense in $O(d)$. Also, G cannot be contained in a hemisphere. A sufficient condition is that the antipodal pairs $\{u \in G : -u \in G\}$ form a generating set of directions for $O(d)$, because for every $u \in \mathbb{S}^{d-1}$ and every $x \in \mathbb{X}$, either $\sigma_u x = \tau_u x$, or $\sigma_u x = \tau_{-u} x$. Must G contain antipodal pairs? Do $d + 1$ directions suffice? (See Fig. 2b.)

Generating sets of directions for $O(d)$ are well understood. For instance, if *(i)* the vectors in G span $\mathbb{R}^d$; *(ii)* G cannot be partitioned into two non-empty mutually orthogonal subsets; and *(iii)* at least one pair of vectors in G encloses an angle that is not a rational multiple of π, then G is a generating set of directions. (The third condition can be relaxed in dimensions $d \geq 3$.) Since d directions $\{u_1, \ldots u_d\}$ in general position are a generating set for $O(d)$, the hypothesis of Theorem 2.2 can be satisfied even by measures whose support has only a finitely many accumulation points.

Theorems 2.1 and 2.2 imply the following statements about Steiner symmetrization.

Corollary 2.3. *Let $\{S_{U_1\dots U_n}\}_{n\geq 1}$ be a sequence of Steiner symmetrizations on $\mathbb{R}^d$ along independently distributed random directions $\{U_i\}$ in $\mathbb{S}^{d-1}$.*

(a) (Convergence of random Steiner symmetrizations). *If*

$$\sum_{i=1}^{\infty} P\big(d(U_i, v_i) < \rho\big) = \infty \tag{2.5}$$

for every radius $\rho > 0$ and every sequence $\{v_i\}$ in $\mathbb{S}^{d-1}$, then

$$P\left(\lim_{n\to\infty} \|S_{U_1\dots U_n} f - f^*\|_\infty = 0 \quad \forall f \in \mathcal{C}_c^+(\mathbb{R}^d)\right) = 1\,. \tag{2.6}$$

(b) (Convergence of i.i.d. Steiner symmetrizations). *The same conclusion holds, if, instead, the random directions $\{U_i\}$ are identically distributed according to a probability measure μ on $\mathbb{S}^{d-1}$ whose support contains a generating set of directions for $O(d)$.*

3. Related work and outline of the proofs

The literature contains several different constructions for convergent sequences of rearrangements. In their proof of the isoperimetric inequality, Carathéodory and Study recursively choose the direction u_n of the next Steiner symmetrization such that $A_n = S_{u_n} A_{n-1}$ is as close to the ball as possible [10]. Lyusternik proposed a sequence that alternates Steiner symmetrization in the d-th coordinate direction with Schwarz symmetrization in the complementary coordinate hyperplane and a well-chosen rotation [21]. Brascamp, Lieb, and Luttinger alternate Steiner symmetrization in all coordinate directions with a rotation [7]. The constructions of Lyusternik and Brascamp-Lieb-Luttinger yield *universal* sequences, which work for all nonnegative functions on $\mathbb{R}^d$ that vanish at infinity.

A number of authors have addressed the question of what distinguishes convergent sequences of Steiner symmetrizations, and how to describe their limits. Eggleston proved that full rotational symmetry can be achieved by iterating Steiner symmetrization in d directions that satisfy a non-degeneracy condition [15, p. 98f]. Klain recently showed that iterating any finite set of Steiner symmetrizations on a convex body results in a limiting body that is symmetric under the subgroup of $O(d)$ generated by the corresponding reflections [18]. On the other hand, Steiner symmetrizations along a dense set of directions may or may not converge to the symmetric decreasing rearrangement, depending on the order in which they are executed [6]. We note in passing that, although the last three results are stated for convex sets, the proofs are readily adapted to functions in $\mathcal{C}_c^+(\mathbb{R}^d)$, with the Arzelà-Ascoli theorem providing the requisite compactness in place of the Blaschke selection theorem. By choosing the measure in Corollary 2.3b to be supported on a finite generating set of directions, we obtain an analogue of Eggleston's theorem for random sequences.

Finding even one convergent sequence of polarizations is more difficult, because it is not enough to iterate a finite collection of polarizations. Baernstein-Taylor,

Benyamini, and Brock-Solynin argue by compactness that the set of functions that can be reached by some finite number of polarizations from a function f contains f^* in its closure [2, 4, 8]. The greedy strategy of Carathéodory and Study also works for the case of polarizations. Both constructions result in sequences that depend on the initial function. A universal sequence was produced by van Schaftingen [28].

In these papers, considerable effort goes into the construction of convergent (or non-convergent) sequences that are rather special. The question whether a randomly chosen sequence converges with probability one was first raised by Mani-Levitska [22]. He conjectured that for compact subsets of $\mathbb{R}^d$, a sequence of Steiner symmetrizations in directions chosen uniformly at random should converge in Hausdorff distance to the ball of the same volume, and verified this for convex sets.

The Mani-Levitska conjecture was settled by van Schaftingen for a larger class of rearrangements that have the same monotonicity, volume-preserving, and smoothing properties as the symmetric decreasing rearrangement [27]. We paraphrase his results for the case of polarization. Van Schaftingen proves the convergence statement in Eq. (2.2) under the assumption that the random variables W_i are independent and their distribution satisfy the uniform bound

$$\liminf_{n\to\infty} P(d(\sigma_{W_n} a, b) < \rho) > 0 \tag{3.1}$$

for every $a, b \in \mathbb{X}$ and every $\rho > 0$. In the proof, he first constructs a *universal* sequence, that is, a single non-random sequence $\{\omega_i\}_{i\geq 1}$ in Ω such that the symmetrizations $S_{\omega_1\dots\omega_n} f$ converge uniformly to f^* for every $f \in \mathcal{C}_c^+(\mathbb{X})$. Eq. (3.1) implies that typical random sequences closely follow the universal sequence for arbitrarily long finite segments, i.e., for every $\rho > 0$ and every integer $N \geq 1$,

$$P(\exists k : d(W_{k+n}, \omega_n) < \rho \text{ for } n = 1, \dots, N) = 1\,.$$

After taking a countable intersection over $\rho_N = \frac{1}{N}$ and $N \in \mathbb{N}$, Eq. (2.2) follows with a continuity argument.

The condition in Eq. (3.1) is stronger than the corresponding assumption of Theorem 2.1. To see this, let $\{a_i\}$, $\{b_i\}$ be a pair of bounded sequences in $\mathbb{X}$, and choose a pair of subsequences $\{a_{i_k}\}$, $\{b_{i_k}\}$ that converge to limits a and b. For k sufficiently large,

$$d(\sigma_\omega a, b) < \frac{\rho}{2} \ \Rightarrow\ d(\sigma_\omega a_{i_k}, b_{i_k}) < \rho\,.$$

If Eq. (3.1) holds, then $P(d(\sigma_{W_{i_k}} a_{i_k}, b_{i_k}) < \rho)$ does not converge to zero, and the series in Eq. (2.1) diverges. We later show examples that satisfy Eq. (2.1) but not Eq. (3.1).

Independently, Volčič has given a direct geometric proof for the convergence of Steiner symmetrizations along uniformly distributed random directions [29]. His proof is phrased as a Borel-Cantelli estimate, which suggests that pairwise independence of the W_i might suffice for convergence (see [14, p. 50-51]). Upon closer inspection, there is a conditioning argument where the independence of the W_i comes into play. It is an open question if convergence can be proved under weaker independence assumptions.

We are not aware of any prior work on rates of convergence for polarizations. There are, however, some very nice results regarding rates of convergence for Steiner symmetrizations of convex bodies. Klartag proved that for every convex body $K \subset \mathbb{R}^d$ and every $0 < \varepsilon < 1/2$, there exists a sequence of $n = \lceil cd^4 \log^2(1/\varepsilon) \rceil$ Steiner symmetrizations $u_1, \ldots, u_n$ such that

$$d_H(\partial S_{u_1 \ldots u_n} K, \partial K^*) \leq \varepsilon \cdot \text{radius}\,(K^*)\,, \tag{3.2}$$

in other words, $(1-\varepsilon)K^* \subset S_{u_1 \ldots u_n} K \subset (1+\varepsilon)K^*$. This means that the distance from a ball decays faster than every polynomial [19, Theorem 1.5]. Remarkably, c is a numerical constant that depends neither on K nor on the dimension. The control over the dimension builds on the earlier result of Klartag and Milman [20] that $3d$ Steiner symmetrizations suffice to reduce the ratio between outradius and inradius of a convex set to a numerical constant. Around the same time, Bianchi and Gronchi established bounds on the rate of convergence in the other direction [5]. For each n and every dimension d, they construct centrally symmetric convex bodies in $\mathbb{R}^d$ whose Hausdorff distance from a ball cannot be decreased by *any* sequence of n successive Steiner symmetrizations. Their construction yields a lower bound on the distance from a ball for arbitrary infinite sequences of Steiner symmetrizations. Klartag's results have recently been extended to random symmetrizations of convex bodies [12]. It is not known whether convergence is in fact exponential, and whether Klartag's convergence estimates can be generalized to non-convex sets.

The proofs of Mani-Levitska, van Schaftingen, and Volčič involve a detailed analysis of typical sample paths. Since they rely on compactness and density arguments, they do not yield bounds on the rate of convergence. In contrast, Bianchi-Gronchi and Klartag use probabilistic methods to find non-random sequences with desired properties. The construction of Bianchi and Gronchi takes advantage of ergodic properties of reflections. Klartag views the rearrangement composed of a random rotation followed by Steiner symmetrizations in each of the d coordinate directions as one step of a Markov chain on convex bodies. He replaces the Steiner symmetrizations by Minkowski symmetrizations to obtain a simpler Markov chain, which acts on the support function of a convex body as a random orthogonal projection in L_2. Since this simpler process is a strict contraction on the spherical harmonics of each positive order, the support function converges exponentially (in expected L_2-distance) to a constant. He finally obtains Eq. (3.2) from a subtle geometric comparison argument.

We combine an analytical approach similar to Klartag's with the geometric techniques used by Volčič. The sequence $\{S_{W_1 \ldots W_n} f\}_{n \geq 1}$ defines a Markov chain on the space $\mathcal{C}_c^+(\mathbb{X})$. We use that the functional

$$\mathcal{I}(f) = \int_{\mathbb{X}} f(x)\, d(x, o)\, dm(x) \tag{3.3}$$

decreases under each polarization, and make Volčič's conditioning argument explicit by appealing to the Markov property. Here, $dm(x)$ denotes integration with respect to the standard Riemannian volume on $\mathbb{X} = \mathbb{S}^d$, $\mathbb{R}^d$, or $\mathbb{H}^d$. For the proof of Theorem 2.1, we quantify the expected value of the drop $\mathcal{I}(f) - \mathcal{I}(S_W f)$ in terms

of $||f - f^*||_\infty$ and the modulus of continuity of f. Since the expected drop goes to zero, $S_{W_1 \dots W_n} f$ converges uniformly to f^*.

For the case of i.i.d. polarizations considered in Theorem 2.2, the challenge is that their distribution may be supported on a small set. Here, we resort to a compactness argument. By monotonicity, $\mathcal{I}(S_{W_1 \dots W_n} f)$ approaches a limiting value. Under the assumptions of the theorem, the drop of $\mathcal{I}$ has *strictly positive expectation* unless $f = f^*$ (Lemma 4.3). This forces the limits of convergent subsequences to be invariant under a family of transformations (the folding maps τ_u parametrized by Eq. (2.4)), which play the role of *competing symmetries* [11]: the only functions that are invariant under the entire family are constant on each centered sphere.

Our estimates for the expected drop of $\mathcal{I}$ imply bounds on the rate of convergence that depend on the modulus of continuity of f and the distribution of the W_i. In the case where the W_i are uniformly distributed on a suitable subset of Ω, we show that there exists a numerical constant c such that

$$E(||S_{W_1 \dots W_n} f - f^*||_\infty) \leq cL \operatorname{Lip}(f)\, n^{-\frac{1}{d+1}}$$

for every Lipschitz continuous nonnegative function f on $\mathbb{R}^d$ with support in B_L (Proposition 5.2). On the other hand, there exist Lipschitz continuous functions f with support in B_L such that

$$E(||S_{W_1 \dots W_n} f - f^*||_\infty) \geq c\, ||f - f^*||_\infty\, q^n\,,$$

where $c > 0$ and $q \in (0,1)$ are numerical constants (Proposition 6.1a).

For Steiner symmetrization, we use that

$$\mathcal{I}(f^*) \;\leq\; \mathcal{I}(S_u f) \;\leq\; \mathcal{I}(S_{(r,\pm u)} f) \;\leq\; \mathcal{I}(f) \tag{3.4}$$

for every $u \in \mathbb{S}^{d-1}$ and all $r > 0$ to bound the expected value of the drop $\mathcal{I}(f) - \mathcal{I}(S_U f)$ under a random Steiner symmetrization from below by the corresponding estimate for a random polarization (Corollary 2.3). By the same token, the power-law bounds on the rate of convergence extend to Steiner symmetrizations along uniformly distributed directions (Corollary 5.4). Since we ignore that Steiner symmetrization reduces perimeter, these bounds cannot be sharp, but to our knowledge they are the only available bounds that do not require convexity. It is an open question whether the sequence converges exponentially, and how the rate of convergence depends on the dimension. Is it more effective to alternate Steiner symmetrizations along the coordinate directions with a random rotation, as in [19]? Does it help to adapt the sequence to the function? Do polarizations converge more slowly, perhaps following a power law?

4. Almost sure convergence

We start by preparing some tools for the proof of the main results. Let $\mathcal{I}$ be the functional defined in Eq. (3.3). The first lemma is a well-known identity, which is related to the Hardy-Littlewood inequality $\int fg \leq \int f^* g^*$. We reproduce its proof here for the convenience of the reader.

Lemma 4.1 (Polarization identity)**.** *Let f be a nonnegative measurable function with $\mathcal{I}(f) < \infty$, and let S_ω be a polarization. Then*

$$\mathcal{I}(f) - \mathcal{I}(S_\omega f) = \int_{\mathbb{X}} [f(\sigma_\omega x) - f(x)]^+ \, [d(\sigma_\omega x, o) - d(x, o)]^+ \, dm(x)\,.$$

In particular, $\mathcal{I}(f) > \mathcal{I}(S_\omega f)$ unless $S_\omega f = f$ almost everywhere.

Proof. We rewrite the functional as an integral over the positive half-space H^+ associated with ω,

$$\mathcal{I}(f) - \mathcal{I}(S_\omega f) = \int_{H^+} \big\{ (f(x) - S_\omega f(x)) d(x, o) + (f(\bar{x}) - S_\omega f(\bar{x})) d(\bar{x}, o) \big\} \, dm(x)\,,$$

where $\bar{x} = \sigma_\omega(x)$. If $f(x) \geq f(\bar{x})$ for some $x \in H^+$, then the values of $S_\omega f$ at x and $\bar{x}$ agree with the corresponding values of f, and the integrand vanishes at x. If, on the other hand $f(x) < f(\bar{x})$, then the values are swapped for $S_\omega f$, and the integrand becomes $(f(x) - f(\bar{x}))(d(x, o) - d(\bar{x}, o))$, where both factors are negative. We switch the signs, collect terms, and integrate to obtain the claim. $\square$

The next lemma is the key ingredient in the proof of Theorem 2.1.

Lemma 4.2 (Expected drop of $\mathcal{I}$)**.** *Let f be a nonnegative continuous function with compact support in $B_L \subset \mathbb{X}$ for some $L > 0$ and modulus of continuity η. Set $\varepsilon = ||f - f^*||_\infty$, let $\rho > 0$ be so small that $\eta(\rho) \leq \frac{\varepsilon}{8}$, and let $W = (R, U)$ be a random variable on Ω, as described above. Then*

$$E(\mathcal{I}(f) - \mathcal{I}(S_W f)) \geq C_\varepsilon \cdot \inf_{x,b} P(d(\sigma_W x, b) < \rho)\,, \tag{4.1}$$

where $C_\varepsilon = \varepsilon \rho \, m(B_\rho)/2$, and the infimum extends over x, b with $d(x, o) + 2\rho \leq d(b, o) \leq L - \rho$. Furthermore, on $\mathbb{X} = \mathbb{R}^d$,

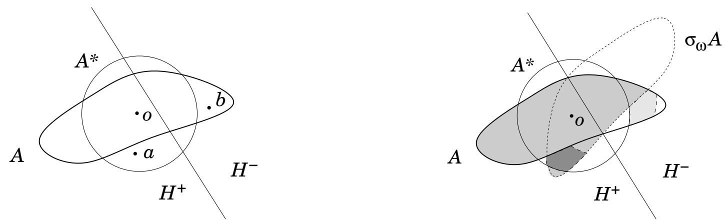
$$E(\mathcal{I}(f) - \mathcal{I}(S_U f)) \geq C'_\varepsilon \cdot \inf_{v \in \mathbb{S}^{d-1}} P(2L \sin d(U, v) < \rho)\,, \tag{4.2}$$

where $C'_\varepsilon = \varepsilon \rho \, m(B_\rho)/8$.

Proof. We first construct a pair of points $a, b \in \mathbb{X}$ such that

$$d(b, o) \geq d(a, o) + 4\rho\,, \quad f(b) \geq f(a) + \frac{\varepsilon}{2}$$

(see Fig. 3a). By assumption, there exists a point x_0 with $|f(x_0) - f^*(x_0)| = \varepsilon$. Set $t = \frac{1}{2}(f(x_0) + f^*(x_0))$, let $A = \{x : f(x) > t\}$, and let A^* be the corresponding level set of f^*. If $f(x_0) < f^*(x_0)$, we set $a = x_0$. By construction, $a \in A^* \setminus A$. Since this set is open and non-empty, it has positive volume, and therefore $A \setminus A^*$, having the same volume, is non-empty. Let $b \in A \setminus A^*$. Then $f^*(a) - f^*(b) > f^*(a) - t = \varepsilon/2$. Similarly, if $f(x_0) > f^*(x_0)$, we set $b = x_0 \in A \setminus A^*$, find $a \in A^* \setminus A$, and note that $f^*(a) - f^*(b) > t - f^*(b) = \varepsilon/2$. Since the modulus of continuity of f is valid also for f^* and $\eta(4\rho) \leq \varepsilon/2$, we have $d(b, o) - d(a, o) \geq 4\rho$.



(a) Choice of a, b in Lemma 4.2 (b) The estimate in Eq. (5.3)

FIGURE 3. Polarization swaps the part of $A \setminus \sigma_\omega A$ that lies in in H^- with its mirror image in H^+. If A is a level set of f and σ_ω is the reflection that maps a to b, then $\mathcal{I}(S_\omega f) < \mathcal{I}(f)$. The volume of $A \triangle A^*$ decreases by the combined volume of the light and dark shaded subsets.

By Lemma 4.1 and Fubini's theorem, a random polarization S_W satisfies

$$\begin{aligned} & E(\mathcal{I}(f) - \mathcal{I}(S_W f)) \\ & \qquad = \quad E\left(\int_{\mathbb{X}} [f(\sigma_W x) - f(x)]^+ \, [d(\sigma_W x, o) - d(x,o)]^+ \, dm(x) \right) \\ & \qquad \geq \quad \frac{\varepsilon\rho}{2} \int_{B_\rho(a)} P(d(\sigma_W x, b) < \rho) \, dm(x) \,, \end{aligned}$$

because the choice of a and b ensures that $f(\sigma_W x) - f(x) \geq \varepsilon/4$ and $d(\sigma_W x, o) - d(x, o) \geq 2\rho$ for all $x \in B_\rho(a)$ with $d(\sigma_W x, b) < \rho$. Eq. (4.1) follows by minimizing over x, a and b and evaluating the integral.

For a random Steiner symmetrization S_U, we use Eq. (3.4) to obtain

$$\begin{aligned} & E(\mathcal{I}(f) - \mathcal{I}(S_U f)) \\ & \qquad \geq \quad E\left(\sup_{r>0,\pm} \int_{\mathbb{X}} [f(\sigma_{(r,\pm U)} x) - f(x)]^+ \left[|\sigma_{(r,\pm U)} x| - |x| \right]^+ dm(x) \right) \\ & \qquad \geq \quad E\left(\mathbb{1}_{\inf |\sigma_{(r,\pm U)} a - b| < \rho} \int_{B_\rho(a)} \frac{\varepsilon\rho}{4} \, dm(x) \right) \\ & \qquad \geq \quad \frac{\varepsilon\rho}{8} \, m(B_\rho(a)) \cdot P(2L \sin d(U, v) < \rho) \,, \end{aligned}$$

where v is the unit vector in the direction of $b - a$, and $d(U, v)$ is the enclosed angle. In the second line, the infimum runs over $(r > 0, \pm)$, and we have used that $f(\sigma x) - f(x) \geq \varepsilon/8$ and $|\sigma x| - |a| > \rho$ whenever $|x - a| < \rho$ and $|\sigma a - b| < \rho$. In the last line, we have estimated the infimum by

$$\inf_{r>0,\pm} |\sigma_{(r,\pm U)} a - b| \;=\; \inf_{t \in \mathbb{R}} |a + tU - b| \;\leq\; (|a| + |b|) \sin d(U, v) \,,$$

and applied Fubini's theorem. □

Proof of Theorem 2.1. Given $f \in \mathcal{C}_c^+(\mathbb{R}^d)$, let $F_n = S_{W_1 \ldots W_n} f$ be the result of n random polarizations of f. Since $F_n = S_{W_n} F_{n-1}$, the sequence $\mathcal{I}(F_n)$ decreases

monotonically and satisfies

$$\mathcal{I}(f) \ \geq\ \mathcal{I}(F_{n-1}) \ \geq\ \mathcal{I}(F_n) \ \geq\ \mathcal{I}(f^*)\,.$$

By writing the difference as a telescoping sum and taking expectations, this implies that

$$\begin{aligned}
&\mathcal{I}(f) - \mathcal{I}(f^*) \\
&\quad \geq \ E\left(\sum_{n=1}^{\infty} \mathcal{I}(F_{n-1}) - \mathcal{I}(S_{W_n}F_{n-1})\right) \\
&\quad = \ \sum_{n=1}^{\infty} E(E(\mathcal{I}(F_{n-1}) - \mathcal{I}(S_{W_n}F_{n-1}) \mid W_1 \ldots W_{n-1})) \\
&\quad \geq \ C_\varepsilon \cdot \sum_{n=1}^{\infty} \left\{ \inf_{x,b} P(d(\sigma_{W_n}x, b) > \rho) \cdot P(||F_{n-1} - f^*||_\infty \geq \varepsilon) \right\}, \qquad (4.3)
\end{aligned}$$

where the infimum extends over all x, b with $d(x,o) + 2\rho \leq d(b,o) \leq L - \rho$, and C_ε, ρ, and L are positive constants that depend on f. We have used the Markov property in the second step, and applied Eq. (4.1) of Lemma 4.2 in the third. In particular, the sum in Eq. (4.3) converges. Since the first factors in the product are not summable by Eq. (2.1), the second factors must have zero as an accumulation point. By monotonicity, they converge to zero. Since $\varepsilon > 0$ was arbitrary, we conclude that

$$P\Big(\lim_{n\to\infty} ||F_n - f^*||_\infty = 0\Big) = 1\,.$$

This establishes Eq. (1.1).

To complete the proof, we choose a countable dense subset $\mathcal{G} \subset \mathcal{C}_c^+$. Let $\{\omega_i\}_{i\geq 1}$ be a sequence in Ω. Since polarizations and the symmetric decreasing rearrangement contract uniform distances, we have for every pair of functions $f, g \in \mathcal{C}_c^+$ and every $n \geq 1$,

$$||S_{\omega_1\ldots\omega_n} f - f^*||_\infty \leq 2||f-g||_\infty + ||S_{\omega_1\ldots\omega_n} g - g^*||_\infty\,.$$

We take $n \to \infty$ and minimize over $g \in \mathcal{G}$ to obtain, by the density of $\mathcal{G}$,

$$\begin{aligned}
\lim_{n\to\infty} ||S_{\omega_1\ldots\omega_n} f - f^*||_\infty &\leq \inf_{g\in\mathcal{G}} \Big\{ 2||f-g||_\infty + \lim_{n\to\infty} ||S_{\omega_1\ldots\omega_n} g - g^*||_\infty \Big\} \\
&\leq \sup_{g\in\mathcal{G}} \ \lim_{n\to\infty} ||S_{\omega_1\ldots\omega_n} g - g^*||_\infty\,.
\end{aligned}$$

Since $\mathcal{G}$ is countable, it follows that

$$\begin{aligned}
&P\left(\exists f \in \mathcal{C}_c^+ : \lim_{n\to\infty} ||S_{W_1\ldots W_n} f - f^*||_\infty > 0\right) \\
&\qquad\qquad \leq \ P\left(\exists g \in \mathcal{G} : \lim_{n\to\infty} ||S_{W_1\ldots W_n} g - g^*||_\infty > 0\right) \\
&\qquad\qquad \leq \ \sum_{g\in\mathcal{G}} P\left(\lim_{n\to\infty} ||S_{W_1\ldots W_n} g - g^*||_\infty > 0\right) \\
&\qquad\qquad = \ 0\,,
\end{aligned}$$

proving Eq. (2.2). ∎

For the proof of Theorem 2.2, we need one more lemma.

Lemma 4.3 (Identification of symmetric decreasing functions)**.** *Let $f \in \mathcal{C}_c^+(\mathbb{X})$.*

(a) (by polarization). *Let W be a random variable on Ω whose distribution satisfies $\mu(R = 0) = 0$. If the orbit of each $x \in \mathbb{S}^{d-1}$ under $G = \{u \in \mathbb{S}^{d-1} : (0, u) \in \operatorname{support} \mu\}$ is dense in $\mathbb{S}^{d-1}$, then*

$$E(\mathcal{I}(S_W f)) = \mathcal{I}(f) \quad \Longleftrightarrow \quad f = f^* \,.$$

(b) (by Steiner symmetrization). *Let U be a random variable on $\mathbb{S}^{d-1}$, and let μ be its probability distribution. If the support of μ contains a generating set of directions for $O(d)$, then*

$$E(\mathcal{I}(S_U f)) = \mathcal{I}(f) \quad \Longleftrightarrow \quad f = f^* \,.$$

Proof. For part (a), suppose that $E(\mathcal{I}(S_W f)) = \mathcal{I}(f)$. It follows from Lemma 4.1 that $\mathcal{I}(S_\omega f) = \mathcal{I}(f)$, and hence $S_\omega f = f$, for μ-a.e. ω. This means that $f(\tau_\omega x) \geq f(x)$ for μ-a.e. ω and all $x \in \mathbb{X}$.

Let $u \in G$. By assumption, μ assigns strictly positive measure to each neighborhood of $(0, u)$ in Ω. Since $\mu(\{R = 0\}) = 0$, we can find a sequence $\omega_i = (r_i, u_i)$ with $r_i > 0$ that converges to $(0, u)$ such that $f(\tau_{\omega_i} x) \geq f(x)$ for each i and all $x \in \mathbb{X}$. By continuity, $f(\tau_u x) \geq f(x)$ for all $x \in \mathbb{X}$, which means that the value of f increases monotonically along orbits $\tau_{u_n} \dots \tau_{u_1} x$ of G. Since $\mathcal{O}_{G,x}$ is dense in the sphere of radius $|x|$ and f is uniformly continuous, f must be radial.

To see that f is symmetric decreasing, we write it as $f(x) = \phi(d(x, o))$ for some continuous function ϕ. Consider first the cases $\mathbb{X} = \mathbb{R}^d$ and $\mathbb{H}^d$. Given $t > 0$, choose $\omega = (r, u)$ with $0 < r \leq 2t$ such that $f(\tau_\omega x) \geq f(x)$ for all $x \in \mathbb{X}$, and let a be the point with normal coordinates (r, u). The reflection σ_ω maps the centered sphere of radius t to the sphere of the same radius centered at a. Since this sphere contains the points with normal coordinates $(r \pm t, u)$, by the intermediate value theorem it contains for each $s \in (t, t + r]$ a point x with $d(x, o) = s$. Since $d(\sigma_\omega x, o) = t < s$, the point x lies in the negative half-space H_ω^-. It follows that $\phi(s) = f(x) \leq f(\tau_\omega(x)) = \phi(t)$. Iterating the argument, we conclude that $\phi(s) \leq \phi(t)$ for all $s \geq t > 0$. Since ϕ is continuous, we can take $t \to 0$ and conclude that ϕ is non-increasing on $[0, \infty)$. In the case $\mathbb{X} = \mathbb{S}^d$, the above argument remains valid for $t \in (0, \pi)$, provided that $r \leq \min\{2t, \pi - t\}$, and we obtain that ϕ is nonincreasing on $[0, \pi]$. This proves that $f = f^*$.

For part (b), suppose that $E(\mathcal{I}(S_U f)) = \mathcal{I}(f)$. We augment the random direction U to a random variable $W = (R, \pm U)$ on Ω, where R is exponentially distributed on $\mathbb{R}^+$, the positive and negative signs are equally likely, and the three components are independent. Then $E(\mathcal{I}(S_{(R,\pm U)} f) = \mathcal{I}(f)$ by Eq. (3.4). The probability distribution of W is given by the measure $d\nu(r, u) = \frac{1}{2} e^{-r} dr (d\mu(u) + d\mu(-u))$ on Ω. By construction, $\nu(\{R = 0\}) = 0$. Since the support of μ contains a generating set of directions for $O(d)$, the orbit of any vector $x \in \mathbb{S}^{d-1}$ under

$$G = \big\{(0, u) \in \operatorname{support} \nu\big\} = \big\{\pm u : u \in \operatorname{support} \mu\big\}$$

is dense in $\mathbb{S}^{d-1}$. Therefore, ν satisfies the assumptions of part (a), and we conclude that $f = f^*$. Finally, the converse implications hold because $S_\omega f^* = f^*$ for all $\omega \in \Omega$. $\square$

Proof of Theorem 2.2. Let W be a random variable on Ω that is distributed according to the measure μ from the statement of the theorem. Lemma 4.3 guarantees that $E(\mathcal{I}(S_W f)) < \mathcal{I}(f)$ unless $f = f^*$. Let $\mathcal{C}_{L,\eta}$ be the set of all nonnegative continuous functions supported in the ball of radius L whose modulus of continuity is bounded by η. Since $\mathcal{I}$ is continuous in the uniform topology and $\mathcal{C}_{L,\eta}$ is compact by the the Arzelà-Ascoli theorem,

$$h(\varepsilon) := \inf\bigl\{E(\mathcal{I}(f) - \mathcal{I}(S_W f)) : f \in \mathcal{C}_{L,\eta}, ||f - f^*||_\infty \geq \varepsilon\bigr\} > 0$$

for each $\varepsilon > 0$.

Given $f \in \mathcal{C}_c^+$, let η be its modulus of continuity, and assume that f is supported in B_L. Denote by $F_n = S_{W_1 \dots W_n} f$ the result of n random polarizations of f. Since polarization preserves the modulus of continuity and the ball B_L, we have $F_n \in \mathcal{C}_{L,\eta}$. We argue as in the proof of Theorem 2.1 that

$$\begin{aligned} \mathcal{I}(f) - \mathcal{I}(f^*) &\geq \sum_{n=1}^{\infty} E(E(\mathcal{I}(F_{n-1}) - \mathcal{I}(S_{W_n} F_{n-1}) \mid W_1 \dots W_{n-1})) \\ &\geq h(\varepsilon) \cdot \sum_{n=1}^{\infty} P(||F_{n-1} - f^*||_\infty \geq \varepsilon)\,. \end{aligned} \tag{4.4}$$

In the second line, we have used the Markov property and the definition of $h(\varepsilon)$. Since $h(\varepsilon) > 0$, the sequence F_n converges almost surely uniformly to f^*, and Eq. (2.2) follows. $\square$

Proof of Corollary 2.3. We proceed as in the proofs of Theorems 2.1 and 2.2, with Eq. (4.2) and Lemma 4.3b in place of Eq. (4.1) and Lemma 4.3a. $\square$

5. Examples in $\mathbb{R}^d$

The following lemma allows to transform integrals over Ω into integrals over $\mathbb{R}^d$. Geometrically, we map $\omega = (r, u)$ to the image of a point a under the reflection σ_ω. Since for every point $z \neq a$ there exists a unique reflection that maps a to z, this defines a diffeomorphism from $\Omega \setminus \{r = 0\}$ to $\mathbb{R}^d \setminus \{a\}$. For $a = o$, the diffeomorphism agrees with the polar coordinate map.

Lemma 5.1 (Change of variables)**.** *Let $a \in \mathbb{R}^d$. Then*

$$\int_\Omega g(\sigma_\omega a)\, d\omega = \int_{\mathbb{R}^d} g(z)|z - a|^{-(d-1)}\, dz$$

for every measurable function g on $\mathbb{R}^d$ such that the integral on the left hand side converges. Here $d\omega = dr dm(u)$ denotes the uniform measure on Ω.

Proof. Set $z = \sigma_\omega a$. If we write $\omega = (r, u)$ and express $z - a$ in polar coordinates (s, v), then $v = u$ because the lines $x = \xi + tu$ are invariant under σ_ω. If r moves by a certain distance, then z moves by that distance in either the direction of u or in

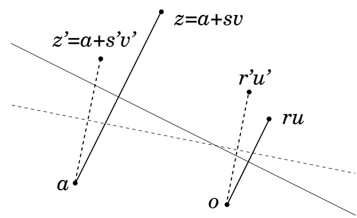


FIGURE 4. The change of variables in Lemma 5.1. In polar coordinates centered at a and o, the volume element transforms as $dsdv = drdu$.

the opposite direction (see Fig. 4). In polar coordinates, the metric on Ω transforms as $(ds)^2 + (dv)^2 = (dr)^2 + (du)^2$. The claim follows by returning to Cartesian coordinates for z. $\square$

We use this formula to construct examples of measures that satisfy the hypothesis of Theorem 2.1 but not Eq. (3.1). Consider the Gaussian probability measure on $\mathbb{R}^d$ whose density is the centered heat kernel at time t. By changing to polar coordinates, we obtain a probability measure on Ω, given by

$$\mu(A) = \frac{1}{(2\pi t)^{\frac{d}{2}}} \int_A e^{-\frac{r^2}{2t}} r^{d-1}\, d\omega\,,$$

where $\omega = (r, u)$. Fix $\rho, L > 0$, let a, b be a pair of points in $\mathbb{R}^d$ with $|a| + 2\rho \leq |b| \leq L - \rho$, and consider the event $\{\omega : d(\sigma_\omega a, b) < \rho\}$. If $z = \sigma_\omega a \in B_\rho(b)$, we use that $|z| - |a| < r < |z| + |a|$ to see that $r \in [2\rho, 2L]$. It follows that there exists a constant C (depending on ρ, L, and the dimension but not on t) such that the density of μ in this region is bounded from below by $Ct^{-\frac{d}{2}} e^{-\frac{2L^2}{t}}$. Changing variables with Lemma 5.1, we estimate

$$\mu(\{\omega : d(\sigma_\omega a, b) < \rho\}) \geq Ct^{-\frac{d}{2}} e^{-\frac{2L^2}{t}} \int_{B_\rho(b)} |z - a|^{-(d-1)}\, dz \ \geq C' t^{-\frac{d}{2}} e^{-\frac{2L^2}{t}}\,.$$

Therefore P, the product of a sequence of such measures, satisfies

$$\sum_{i=1}^{\infty} P(d(\sigma_{W_i} a_i, b_i) < \rho) \geq C' \sum_{i=1}^{\infty} t_i^{-\frac{d}{2}} e^{-\frac{2L^2}{t_i}}$$

for any pair of sequences $\{a_i\}$, $\{b_i\}$ in $\mathbb{R}^d$ with $|a_i| + 2\rho \leq |b_i| \leq L - \rho$. For $t_i = (\log\log i)^{-1}$ the sum diverges as required by Eq. (2.1), but Eq. (3.1) fails because the measures converge weakly to zero. For the sequence $t_i = i^{2/d}$, Eq. (2.1) holds but Eq. (3.1) fails because the measures concentrate on $\{0\} \times \mathbb{S}^{d-1}$.

To give a similar example for Steiner symmetrizations, consider the probability measures on $\mathbb{S}^{d-1}$ defined by the Poisson kernel

$$\mu(A) = \frac{1}{m(\mathbb{S}^{d-1})} \int_A \frac{1 - |z|^2}{|z - u|^d}\, dm(u)\,,$$

where z is a point in the ball, and dm denotes integration with respect to the standard Riemannian volume in $\mathbb{S}^{d-1}$. Since the density of μ with respect to the uniform probability measure on $\mathbb{S}^{d-1}$ is bounded from below by $2^{-(d-1)}(1-|z|)$, the product of such measures satisfies

$$\sum_{i=1}^{\infty} P(d(U_i, v_i) < \rho) \geq 2^{-(d-1)} \frac{m(B_\rho)}{m(\mathbb{S}^{d-1})} \sum_{i=1}^{\infty} (1 - |z_i|)$$

for every sequence $\{v_i\}$ in $\mathbb{S}^{d-1}$. If $z_i = (1 - 1/i)\, u$ for some $u \in \mathbb{S}^{d-1}$, then the sum diverges and Eq. (2.5) holds, but condition (3.1) fails because the measures converge weakly to the point mass at u.

In principle, the proofs of Theorems 2.1 and 2.2 imply weak-type bounds on the rate of convergence. Eq. (4.3) yields that

$$P(||F_n - f^*||_\infty \geq \varepsilon) \leq \frac{\mathcal{I}(f) - \mathcal{I}(f^*)}{C_\varepsilon \cdot \sum_{i=1}^{n} \inf_{x,b} P(d(\sigma_{W_i} x, b) > \rho)},$$

where C_ε and ρ depend on the modulus of continuity of f. Similarly, since $||F_n - f^*||_\infty$ is non-increasing, Eq. (4.4) yields that

$$P(||F_n - f^*||_\infty \geq \varepsilon) \leq \frac{\mathcal{I}(f) - \mathcal{I}(f^*)}{h(\varepsilon)}\, n^{-1},$$

where $h(\varepsilon)$ depends on the distribution of the random polarizations and the modulus of continuity of f. For i.i.d. uniform sequences of rearrangements, we have a more explicit bound:

Proposition 5.2 (Rate of convergence for random polarizations)**.** *If $\{W_i\}_{i\geq 1}$ is a sequence of independent uniformly distributed random variables on $(0, 2L) \times \mathbb{S}^{d-1}$, then*

$$E(||S_{W_1 \ldots W_n} f - f^*||_1) \leq 2d\, m(B_{2L})\, ||f||_\infty\, n^{-1} \tag{5.1}$$

for every nonnegative bounded measurable function on $\mathbb{R}^d$ with support in B_L. If, additionally, f is Hölder continuous with modulus of continuity $\eta(\delta) \leq c\delta^\alpha$ for some $\alpha \in (0, 1]$ and $c > 0$, then

$$E(||S_{W_1 \ldots W_n} f - f^*||_\infty) \leq 10 c L^\alpha n^{-\frac{\alpha}{d+\alpha}}. \tag{5.2}$$

In the proof of the proposition, we will use the following lemma.

Lemma 5.3 (Expected drop in symmetric difference)**.** *If W is a uniformly distributed random variable on $(0, 2L) \times \mathbb{S}^{d-1}$, then*

$$m(A \triangle A^*) - E(m(S_W A \triangle A^*)) \geq \frac{1}{2d\, m(B_{2L})} (m(A \triangle A^*))^2$$

for every measurable set $A \subset B_L$ in $\mathbb{R}^d$.

Proof. Fix $\omega \in \Omega$, and let H^+ and H^- be the half-spaces associated with ω. By construction, polarization swaps the portion of $A \setminus \sigma_\omega A$ that lies in H^- with its mirror image in H^+ (see Fig. 1a). Of these sets, precisely the portion of $A \setminus A^*$

whose reflection lies in $A^* \setminus A$ contributes to the symmetric difference $A \triangle A^*$, twice, see Fig. 3b. But this just means that

$$m(A \triangle A^*) - m(S_\omega A \triangle A^*) = 2m(\{x \in A^* \setminus A : \sigma_\omega(x) \in A \setminus A^*\})\,. \tag{5.3}$$

We compute the expectation, using Fubini's theorem and the change of variables from Lemma 5.1. The result is

$$\begin{aligned} m(A \triangle A^*) &- E(m(S_W A \triangle A^*)) \\ &= 2\int_{A^*\setminus A} P(\sigma_W(x) \in A \setminus A^*)\,dx \\ &= \frac{1}{L\,m(\mathbb{S}^{d-1})} \int_{A^*\setminus A}\int_{A\setminus A^*} |x-z|^{-(d-1)}\,dzdx \\ &\geq \frac{1}{C}(m(A \triangle A^*))^2\,, \end{aligned}$$

where $C = 2d\,m(B_{2L})$. In the last step, we have used that the distance between x and z is at most $2L$, and that A and A^* have the same volume. Note that the Riemannian volume of the unit sphere in $\mathbb{R}^d$ is related to the Lebesgue measure of the unit ball by $m(\mathbb{S}^{d-1}) = d\,m(B_1)$. □

Proof of Proposition 5.2. Consider first the case where $f = \mathbb{1}_A$ for some measurable set $A \subset B_L$, and let $A_n = S_{W_n \dots W_1} A$. By Lemma 5.3, the Markov property, and Jensen's inequality,

$$\begin{aligned} E(m(A_{n-1} \triangle A^*)) &- E(m(A_n \triangle A^*)) \\ &= E\big(m(A_{n-1} \triangle A^*) - E(S_{W_n} A_{n-1} \triangle A^*) \mid W_1, \dots, W_{n-1})\big) \\ &\geq \frac{1}{C} E\big((m(A_{n-1} \triangle A^*))^2\big) \\ &\geq \frac{1}{C}\big(E(m(A_{n-1} \triangle A^*))\big)^2\,, \end{aligned}$$

where $C = 2d\,m(B_{2L})$. This shows that $z_n = C^{-1}E(m(A_n \triangle A^*))$ satisfies the recursion relation $z_n \leq z_{n-1}(1-z_{n-1})$. Since $z_n^{-1} \geq z_{n-1}^{-1} + 1$ and $z_0^{-1} \geq d2^{d+1}$, it follows that

$$E(m(A_n \triangle A^*)) \ \leq C(n + d2^{d+1})^{-1}\,. \tag{5.4}$$

If f is a nonnegative bounded measurable function on B_L, we use the layer-cake principle to write

$$||f - f^*||_1 = \int_0^\infty m(\{f > s\} \triangle \{f^* > s\})\,ds\,,$$

and likewise for $F_n = S_{W_n \dots W_n} f$. Since f is bounded, the integrand vanishes for $s > ||f||_\infty$, and we obtain from Eq. (5.4) that

$$E(||F_n - f^*||_1) \leq C||f||_\infty (n + d2^{d+1})^{-1}\,,$$

proving the first claim.

If f is Hölder continuous, then F_n and and f^* are Hölder continuous with the same modulus of continuity. Let $\varepsilon = ||F_n - f^*||_\infty$, and set $\rho = \eta^{-1}(\varepsilon/4)$. Since

F_n differs from f^* by at least $\varepsilon/2$ on some ball of radius ρ, we have $||F_n - f^*||_{L_1} \geq \varepsilon m(B_\rho)/2$. We obtain from Eq. (5.1) that

$$\begin{aligned} E\big((||F_n - f^*||_\infty)^{1+\frac{d}{\alpha}}\big) &\leq \frac{2(4c)^{\frac{d}{\alpha}}}{m(B_1)} E(||F_n - f^*||_1) \\ &\leq d2^d(4cL^\alpha)^{1+\frac{d}{\alpha}} n^{-1} \,. \end{aligned}$$

Applying Jensen's inequality once more, we arrive at

$$E(||F_n - f^*||_\infty) \leq (d2^d)^{\frac{\alpha}{d+\alpha}} 4cL^\alpha n^{-\frac{\alpha}{d+\alpha}} \,.$$

The leading constant is maximized at $\alpha = 1$ and $d = 6$, and Eq. (5.2) follows. $\square$

By Eq. (3.4), Proposition 5.2 extends to Steiner symmetrization along directions chosen independently and uniformly at random on $\mathbb{S}^{d-1}$.

Corollary 5.4 (Rate of convergence for random Steiner symmetrizations)**.** *If $\{U_i\}_{i\geq 1}$ is a sequence of independent uniformly distributed random variables on $\mathbb{S}^{d-1}$, then*

$$E(||S_{U_1\dots U_n} f - f^*||_1) \leq 2d\, m(B_{2L})\, ||f||_\infty n^{-1}$$

for every nonnegative bounded measurable function with support in B_L. If f is Hölder continuous with modulus of continuity $\eta(\delta) \leq c\delta^\alpha$ for some $\alpha \in (0,1]$ and $c > 0$, then

$$E(||S_{W_1\dots W_n} f - f^*||_\infty) \leq 10cL^\alpha n^{-\frac{\alpha}{d+\alpha}} \,.$$

6. Negative results

In this section, we give some bounds on the rate of convergence that complement Proposition 5.2 and Corollary 5.4, and construct examples where convergence fails. For polarization, we use the function

$$f(x) = [1 - d(x,a)]^+ \,, \tag{6.1}$$

which is supported on $B_1(a)$ and Lipschitz continuous with constant one. Its symmetric decreasing rearrangement is given by $f^*(x) = [1-|x|]^+$, and $||f - f^*||_\infty = \min\{d(a,o), 1\}$. Its polarization at $\omega \in \Omega$ is given by

$$S_\omega f(x) = [1 - d(x, \tau_\omega a)]^+ \,,$$

where τ_ω is the folding map that fixes the positive half-space H_ω^+ and reflects H_ω^- across the separating hyperplane.

Proposition 6.1. *Let f be given by Eq. (6.1).*

(a) (Convergence of random polarizations is not faster than exponential). *If $d(a,o) \leq 1$, then*

$$E(||S_{W_1\dots W_n} f - f^*||_\infty\,) \geq ||f - f^*||_\infty\, 2^{-n}$$

for every sequence $\{W_i\}_{i\geq 1}$ of independent random variables on Ω such that the distribution of each $W_i = (R_i, U_i)$ is symmetric under $U_i \mapsto -U_i$.

(b) (Non-convergence). *If $a \neq o$, then there exists a dense sequence $\{\omega_i\}_{i\geq 1}$ in Ω such that $S_{\omega_1\dots\omega_n} f$ has no limit in $\mathcal{C}_c^+(\mathbb{X})$.*

Proof. A single random polarization results in $S_W f(x) = [1 - d(x, \tau_W a)]^+$. Since $\tau_W a = a$ whenever $a \in H_W^-$, its expected distance from the origin satisfies $E(d(\tau_W a, o)) \geq d(a, o)/2$. By iteration, we have $S_{W_1 \dots W_n} f(x) = [1 - d(x, a_n)]^+$, where $a_n = \tau_{W_n} a_{n-1}$ and $a_0 = a$. By the Markov property, $E(d(a_n, o)) \geq d(a, o) 2^{-n}$, and the first claim follows.

For the second claim, we realize an arbitrary sequence as a subsequence of one for which convergence fails. Given $\{\omega_i\}_{i \geq 1}$, fix $0 < \varepsilon < d(a, o)$ and define $\{\tilde{\omega}_i\}_{i \geq 1}$ as follows. On the odd integers set $\tilde{\omega}_{2n-1} = (\min\{2^{-n}\varepsilon, r_n\}, \pm u_n)$, where $(r_n, u_n) = \omega_n$, and the sign is chosen in such a way that $S_{\tilde{\omega}_1 \dots \tilde{\omega}_{2n-1}} f$ is unchanged by S_{ω_n}. On the even integers, set $\tilde{\omega}_{2n} = \omega_n$. If $\{\omega_i\}$ is dense, then $\{\tilde{\omega}_i\}$ is dense as well.

Set $f_n = S_{\tilde{\omega}_1 \dots \tilde{\omega}_n} f = [1 - d(x, a_n)]^+$. Suppose that f_n converges to some limit g. Then $g(x) = [1 - d(x, b)]^+$ for some b. Let $\omega = (r, u) \in \Omega$ with $r > 0$. By density, we can find a subsequence $\{\tilde{\omega}_{n_k}\}$ that converges to ω. Since both $a_{n_k - 1}$ and $a_{n_k} = \tau_{\tilde{\omega}_{n_k}} a_{n_k - 1}$ converge to b, we must have $\tau_\omega b = b$. Since ω was arbitrary, it follows that $b = o$. On the other hand, $d(b, o) \geq d(a, o) - \varepsilon \sum 2^{-n} > 0$, a contradiction. $\square$

The corresponding bounds for Steiner symmetrizations on $\mathbb{R}^d$ are slightly more involved. As an example, we use the function

$$f(x) = [1 - \langle x, Mx \rangle]^+ \,, \tag{6.2}$$

where M is a positive definite symmetric $d \times d$ matrix. The symmetric decreasing rearrangement of f is $f^*(x) = [1 - \lambda^* |x|^2]^+$, where λ^* is the geometric mean of the eigenvalues of M. The distance from f to f^* satisfies

$$\frac{\lambda_{\max} - \lambda_{\min}}{2\lambda_{\max}} \leq ||f - f^*||_\infty \leq \frac{\lambda_{\max} - \lambda_{\min}}{\lambda_{\min}} \,. \tag{6.3}$$

We will prove the following statements.

Proposition 6.2. *Let f be given by Eq. (6.2) with some positive definite symmetric matrix M.*

(a) (Convergence of random Steiner symmetrizations is not faster than exponential). *If $\{U_i\}$ is a sequence of i.i.d. uniform random variables on $\mathbb{S}^{d-1}$ and the extremal eigenvalues of M satisfy $\lambda_{\max} \leq 2\lambda_{\min}$, then*

$$E(||S_{U_1 \dots U_n} f - f^*||_\infty) \geq \frac{1}{4} ||f - f^*||_\infty \, 3^{-n} \,.$$

(b) (Non-convergence). *If M is not a multiple of the identity, then there exists a dense sequence $\{u_i\}_{i \geq 1}$ in $\mathbb{S}^{d-1}$ such that $S_{u_1 \dots u_n} f$ has no limit in $\mathcal{C}_c^+(\mathbb{R}^d)$.*

We first show that Steiner symmetrization preserves the form of f.

Lemma 6.3 (Steiner symmetrization of ellipsoids)**.** *If f is given by Eq. (6.2), then $S_u f$ has the same form with a positive definite symmetric matrix M' determined by*

$$\langle x, M'x \rangle = \langle x, Mx \rangle - \frac{\langle x, Mu \rangle^2}{\langle u, Mu \rangle} + \langle x, u \rangle^2 \langle u, Mu \rangle \,. \tag{6.4}$$

In particular, u is an eigenvector of M' with eigenvalue $\langle u, Mu \rangle$.

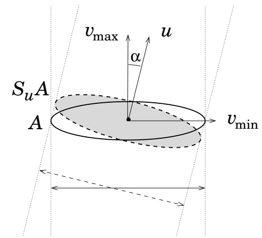


FIGURE 5. Steiner symmetrization of an ellipse. The diameter shrinks at most by a factor $\cos\alpha$.

Proof. Consider a line $x = \xi + tu$ with $\xi \perp u$. The restriction of f to this line, given by

$$t \mapsto [1 - \langle \xi, M\xi\rangle - 2t\langle \xi, Mu\rangle - t^2\langle u, Mu\rangle]^+ ,$$

is symmetric decreasing about $t_0 = -\frac{\langle \xi, Mu\rangle}{\langle u, Mu\rangle}$. By definition, the restriction of $S_u f$ to the line is the symmetrized function

$$t \mapsto [1 - \langle \xi, M\xi\rangle + (t_0^2 - t^2)\langle u, Mu\rangle]^+ ,$$

as required by Eq. (6.4). Since $u^\perp$ and the line through u are invariant subspaces for M', we conclude that u is an eigenvector. The corresponding eigenvalue is $\lambda = \langle u, M'u\rangle = \langle u, Mu\rangle$. $\square$

Remark. An amusing consequence of Lemma 6.3 is that $(d{-}1)$ Steiner symmetrizations suffice to transform an ellipsoid into a ball [20]. To see this let A be an ellipsoid of the same volume as the unit ball. Then $A = \{\langle x, Mx\rangle < 1\} = \{f > 0\}$, where M is a positive definite symmetric matrix of determinant one, and f is given by Eq. (6.2). Set $M_0 = M$, and choose u_1 such that $\langle u_1, Mu_1\rangle = 1$. By Lemma 6.3, $S_{u_1}A = \{\langle x, M_1x\rangle < 1\}$, where M_1 is a positive definite symmetric matrix that has u_1 as an eigenvector with eigenvalue 1. Iteratively choosing u_i orthogonal to $u_1, \dots, u_{i-1}$ such that $\langle u_i, M_{i-1}u_i\rangle = 1$, we arrive at $M_{d-1} = I$, and conclude that $S_{u_{d-1}} \dots S_{u_1} A = A^*$. $\square$

To prove Proposition 6.2, we need to analyze how the extremal eigenvalues of M change under Steiner symmetrization of f. Clearly, their difference decreases, because the inradius of the corresponding ellipsoid grows under Steiner symmetrization, and its outradius shrinks. The following lemma shows that the change in the extremal eigenvalues is small, if the direction of the Steiner symmetrization is either almost parallel or almost orthogonal to the maximizing eigenvector $v_{\max}$ (see Fig. 5).

Lemma 6.4 (Eigenvalue estimate)**.** *Given a symmetric positive definite matrix M with extremal eigenvalues $\lambda_{\max}$, $\lambda_{\min}$ and corresponding normalized eigenvectors $v_{\max}$, $v_{\min}$. Define M' by Eq. (6.4). The extremal eigenvalues $\lambda'_{\max}$, $\lambda'_{\min}$ of M' satisfy*

$$\lambda'_{\max} - \lambda'_{\min} \geq \big(1 - C\psi(\langle u, v_{\max}\rangle) - 2\psi(\langle u, v_{\min}\rangle)\big)\left(\lambda_{\max} - \lambda_{\min}\right), \tag{6.5}$$

where $C = 1 + \lambda_{\max}/\lambda_{\min}$ and $\psi(t) = t^2(1 - t^2)$.

Proof. Let v be a normalized eigenvector of M with eigenvalue λ. From Eq. (6.4), we obtain that

$$\begin{aligned}\langle v, M'v\rangle &= \lambda - \frac{\lambda^2\langle u,v\rangle^2}{\langle u, Mu\rangle} + \langle u,v\rangle^2\langle u, Mu\rangle \\ &= \lambda + \cos^2\!\alpha \sin^2\!\alpha \left(1 + \frac{\lambda}{\langle u, Mu\rangle}\right)(\langle w, Mw\rangle - \lambda)\,.\end{aligned}$$

In the second step, we have expanded $u = \cos\alpha\, v + \sin\alpha\, w$, where w is a unit vector orthogonal to v, and then collected terms. We apply this identity to $v_{\max}$ and use that $\langle u, Mu\rangle \geq \lambda_{\min}$ and $\langle w, Mw\rangle \leq \lambda_{\max}$ to obtain

$$\begin{aligned}\lambda'_{\max} &\geq \langle v_{\max}, M'v_{\max}\rangle \\ &\geq \lambda_{\max} - \left(1 + \frac{\lambda_{\max}}{\lambda_{\min}}\right)\psi(\langle u, v_{\max}\rangle)\,(\lambda_{\max} - \lambda_{\min})\,.\end{aligned}$$

Similarly,

$$\begin{aligned}\lambda'_{\min} &\leq \langle v_{\min}, M'v_{\min}\rangle \\ &\leq \lambda_{\min} + 2\psi(\langle u, v_{\min}\rangle)\,(\lambda_{\max} - \lambda_{\min})\,.\end{aligned}$$

The claim follows by subtracting the two inequalities. □

Lemma 6.5 (Expected change of extremal eigenvalues)**.** *Let f be given by Eq. (6.2) with a positive definite symmetric matrix M whose extremal eigenvalues satisfy $\lambda_{\max} \leq 2\lambda_{\min}$. If U is a uniformly distributed random variable on $\mathbb{S}^{d-1}$, then $S_U f(x) = [1-\langle x, M'x\rangle]^+$, where M' is a positive definite symmetric matrix whose the extremal eigenvalues satisfy*

$$E(\lambda'_{\max} - \lambda'_{\min}) \geq \frac{1}{3}(\lambda_{\max} - \lambda_{\min})\,.$$

Proof. We apply Lemma 6.4 and take expectations. Let v_{max} and $v_{\min}$ be the eigenvectors of M corresponding to $\lambda_{\max}$ and $\lambda_{\min}$, and set $C = 1 + \lambda_{\max}/\lambda_{\min} \leq 3$ and $\psi(t) = t^2(1-t^2)$. By taking advantage of the rotation invariance, we compute $E(\langle U, v\rangle^2) = 1/d$ and $E(\langle U, v\rangle^4) = 3/(d(d+2))$ for all $v \in \mathbb{S}^{d-1}$, see [16, Exercise 63, p. 80]. This results in

$$E(1 - C\psi(\langle U, v_{\max}\rangle) - 2\psi(\langle U, v_{\min}\rangle)) = 1 - (C+2)\left(\frac{1}{d} - \frac{3}{d(d+2)}\right)\,.$$

The claim follows by evaluating the right hand side at $d = 3$, where it assumes its minimum value, and using Eq. (6.5). □

Proof of Proposition 6.2. Let f be given by Eq. (6.2) with some positive definite symmetric matrix M. We first consider the case of a random sequence $F_n = S_{U_1 \dots U_n} f$, where the directions $\{U_i\}$ are independent and uniformly distributed on $\mathbb{S}^{d-1}$. By Lemma 6.3, we can write F_n in the form (6.2) with a positive definite symmetric matrix M_n that is recursively defined by Eq. (6.4) with $u = U_n$. We iterate the estimate in Lemma 6.5, using the Markov property, and obtain that the

gap between the extremal eigenvalues of M_n is at least $(\lambda_{\max}-\lambda_{\min})3^{-n}$. Since we assumed that $\lambda_{\max} \leq 2\lambda_{\min}$, it follows from Eq. (6.3) that

$$\begin{aligned} E(||S_{U_1\dots U_n}f - f^*||_\infty) &\geq \frac{\lambda_{\max}-\lambda_{\min}}{2\lambda_{\max}} 3^{-n} \\ &\geq \frac{1}{4}||f-f^*||_\infty \, 3^{-n} \,. \end{aligned}$$

For the second claim, we proceed as in the proof of Proposition 6.1 by realizing an arbitrary sequence as a subsequence of one for which convergence fails. Given $\{u_i\}_{i\geq 1}$ in $\mathbb{S}^{d-1}$, let $\varepsilon > 0$ so small that $(C+2)\sin^2\varepsilon < 1$, where $C = 1 + \lambda_{\max}/\lambda_{\min}$ as in Lemma 6.5, and construct the sequence $\{v_i\}_{i\geq 1}$ as follows. In the first step, pick v_1 to be a maximizing eigenvector of M. Suppose we have already chosen $v_1, \dots, v_n$ such that $d(v_i, v_{i+1}) \leq \varepsilon/i$ for each $i < n$, and that $u_1, \dots, u_j$ appear as a subsequence. If $d(v_n, u_{j+1}) \leq \varepsilon/n$, pick $v_{n+1} = u_{j+1}$. Otherwise, choose v_{n+1} on the great circle that joints v_n with u_{j+1} in such a way that $d(v_n, v_{n+1}) = \varepsilon/n$ and $d(v_{n+1}, u_{j+1}) = d(v_n, u_{j+1}) - \varepsilon/n$. Since $\sum \varepsilon/n$ diverges, the entire sequence $\{u_i\}$ is incorporated as a subsequence into $\{v_i\}$. If $\{u_i\}$ is dense, so is $\{v_i\}$.

Let $f_n = S_{v_1\dots v_n}f = [1 - \langle x, M_n x\rangle]^+$. If f_n converges to some limit g, then g is given by Eq. (6.2) with some positive definite symmetric matrix N of the same determinant as M. Since $\{v_i\}$ is dense in $\mathbb{S}^{d-1}$, we find that N is necessarily a multiple of the identity because g is invariant under every Steiner symmetrization. On the other hand, we can estimate the extremal eigenvalues of N as follows. By construction, v_n is an eigenvector of M_n. Since $d(v_n, v_{n+1}) \leq \varepsilon/n$ and the other eigenvectors of M_n are orthogonal to v_n, we have that $\psi(\langle v_{n+1}, v\rangle) \leq \sin^2(\varepsilon/n)$ for each eigenvector v of M_n. Iterating Lemma 6.4, we see that the gap between the extremal eigenvalues of N is at least $(\lambda_{\max}-\lambda_{\min})\prod(1-(C+2)\sin^2(\varepsilon/n)) > 0$, a contradiction. □

7. Compact sets

We finally collect the implication of our results for compact sets. Under the assumptions of Theorems 2.1 and 2.2, random polarizations of functions in L_p^+ also converge almost surely in L_p, see Eq. (2.3). In particular, for $p = 1$,

$$P\Big(\lim_{n\to\infty} m(S_{W_1\dots W_n}A \,\triangle\, A^*) = 0 \quad \forall A \subset \mathbb{X} \text{ with } m(A) < \infty\Big) = 1 \,. \tag{7.1}$$

This is another equivalent restatement of Eq. (2.2). We now establish the corresponding convergence result for the Hausdorff distance.

The topology defined by the Hausdorff metric on the space of compact sets is not comparable to the topology of symmetric difference. Moreover, polarization is not continuous with respect to Hausdorff distance. To give a simple example, consider a reflection σ that does not fix the origin, and let a be the image of the origin under σ. By definition, $S(\{o, a\}) = \{o, a\}$. Let $\{a_i\}_{i\geq 1}$ be a sequence in $\mathbb{X}$ with $a_i \neq a$ for all i that converges to a. The sequence of two-point sets $\{o, a_i\}$ clearly converges to $\{o, a\}$. Since the reflected sequence $\{\sigma a_i\}$ converges to the origin, we have that

$d(a_i, o) > d(\sigma a_i, o)$ and therefore $a_i \in H^-$ for i large enough. It follows that $S(\{o, a_i\}) = \{o, \sigma a_i\}$, which converges in Hausdorff distance to $\{o\}$.

Nevertheless, convergence of a sequence of polarizations in Hausdorff distance to a ball implies convergence in symmetric difference. To see this, let K be a compact set of positive volume, and consider a sequence $K_n = S_{\omega_1 \dots \omega_n} K$. If K_n converges to (the closure of) K^* in Hausdorff distance, then the radius of the smallest centered ball containing K converges to the radius of K^*, which implies that $m(K_n \setminus K^*)$ converges to zero. Since K_n and K^* have the same volume, $m(K^* \setminus K_n)$ goes to zero as well.

In the other direction, we can obtain convergence in Hausdorff distance from the uniform convergence statement in Eq. (2.2) by realizing a given compact set as a level set of a continuous function.

Proposition 7.1 (Convergence in Hausdorff distance). *If a random sequence $\{W_i\}$ satisfies the assumptions of either Theorem 2.1 or Theorem 2.2, then*

$$P\left(\lim_{n\to\infty} d_H\big(S_{W_1\dots W_n}K, K^*\big) = 0 \quad \forall \text{ compact } K \subset \mathbb{X} \text{ with } m(K) > 0\right) = 1\,,$$

and

$$P\left(\lim_{n\to\infty} d_H\big(\partial S_{W_1\dots W_n}K, \partial K^*\big) = 0 \quad \begin{array}{l}\forall \text{ compact } K \subset \mathbb{X} \\ \text{with } m(K) > 0 \text{ and } m(\partial K) = 0\end{array}\right) = 1\,.$$

Proof. Set $K_n = S_{W_1\dots W_n}K$. We consider the two pieces of the Hausdorff distance from K_n to K^* separately. If $\operatorname{dist}(x, K_n) = \delta > 0$ for some $x \in K^*$, then $m(K_n \triangle K^*) \geq 2m(B_\delta(x) \cap K^*) > 0$. Therefore Eq. (7.1) implies that

$$\sup_{x\in K^*} \operatorname{dist}(x, K_n) \to 0 \quad (n \to \infty) \quad \text{almost surely}$$

simultaneously for all K.

To control the other piece of $d_H(K_n, K^*)$, we use the auxiliary function $f(x) = [1 - \operatorname{dist}(x, K)]^+$. By definition, the level set of f at height $1-t$ is the outer parallel set $\{x : \operatorname{dist}(x, K) < t\}$. The level set of f^* at that height is the centered ball of the same volume. Its radius $\rho(t)$, defined by

$$B_{\rho(t)} = \{x : \operatorname{dist}(x, K) < t\}^* \quad (t > 0)$$

depends continuously on t and converges to the radius of K^* as $t \to 0$. Set $F_n = S_{W_1\dots W_n} f$. Since $F_n(x) = 1$ for all $x \in K_n$,

$$\begin{aligned} \sup_{x\in K_n} \operatorname{dist}(x, K^*) &= \sup_{x\in K_n\setminus K^*} \rho(F_n(x) - f^*(x)) - \text{radius}\,(K^*) \\ &\to 0 \quad (n \to \infty) \quad \text{almost surely} \end{aligned} \tag{7.2}$$

by Theorem 2.1. This proves the first claim.

If ∂K has zero volume, we continuously extend the function ρ such that

$$B_{\rho(0)} = K^*\,, \qquad B_{\rho(t)} = \{x : \operatorname{dist}(x, \mathbb{X} \setminus K) > -t\}^* \quad (t < 0)\,,$$

and replace the auxiliary function with

$$f(x) = \big[h + \operatorname{dist}(x, \mathbb{X} \setminus K) - \operatorname{dist}(x, K)\big]^+\,, \tag{7.3}$$

where $h > 0$ is an arbitrary constant. The level sets of f at heights below h are outer parallel sets of K, while the level sets at heights above h are inner parallel sets. It follows that

$$\begin{aligned} d_H(\partial K_n, \partial K^*) &= \sup_{x \in \partial K_n} |\rho(h - f^*(x)) - \text{radius}\,(K^*)| \\ &\le \max_{\pm} |\rho(\pm||F_n - f^*||_\infty) - \rho(0)| \\ &\to 0 \quad (n \to \infty) \quad \text{almost surely.} \end{aligned} \tag{7.4}$$

In the second line, we have used that $F_n = h$ on ∂K_n. The last line follows from Theorem 2.1 and the continuity of ρ. □

Similar arguments can be used to bound the rate of convergence for sets with additional regularity properties. Let K be a compact set in $\mathbb{R}^d$, and define f and ρ as in the proof of the second claim of Proposition 7.1. Assume that $K \subset B_L$, and that ρ is differentiable at $t = 0$ with $\rho'(0) = \text{Per}\,(K)/\text{Per}\,(K^*)$. By Proposition 5.2 there exists a sequence $\{\omega_i\}$ such that

$$||S_{\omega_1 \ldots \omega_n} f - f^*||_\infty \le 10(L + h)\, n^{-\frac{1}{d+1}} .$$

Expanding ρ about $t = 0$, we obtain from Eq. (7.4) that

$$\begin{aligned} d_H(\partial K_n, \partial K^*) &\le \rho'(0)\,(1 + o(1))\,||S_{\omega_n \ldots \omega_1} f - f^*||_\infty \\ &\le C \cdot \text{radius}\,(K^*) \cdot \frac{\text{Per}\,(K)}{\text{Per}\,(K^*)}\,(1 + o(1))\, n^{-\frac{1}{d+1}} \end{aligned} \tag{7.5}$$

as $n \to \infty$, where $C = 10(L + h)/\rho(0)$. After dropping an initial segment $n \le N$ from the sequence, we may replace L with the radius of the smallest centered ball containing K_N. Choosing N sufficiently large and h sufficiently small, we can find a sequence of polarizations where Eq. (7.5) holds with $C = 10$.

Remark. The conclusions of Proposition 7.1 also hold for random Steiner symmetrizations that satisfy the assumptions of Corollary 2.3. Likewise, Eq. (7.5) applies to sequences of Steiner symmetrizations along i.i.d. uniformly distributed directions. However, in view of Klartag's result for convex sets, we expect such sequences to converge more rapidly (see Eq. (3.2)).

Acknowledgments. This paper is based on results from M.F.'s 2010 Master's thesis at the University of Toronto [17]. Our research was supported in part by NSERC through Discovery Grant No. 311685-10 (Burchard) and an Alexander Graham Bell Canada Graduate Scholarship (Fortier). A.B. wishes to thank Gerhard Huisken (Albert-Einstein Institut in Golm), Bernold Fiedler (Freie Universität Berlin), Nicola Fusco (Università di Napoli Federico II) and Adèle Ferone (Seconda Università di Napoli, Caserta) for hospitality during a sabbatical in 2008/09. Special thanks go to Aljoša Volčič for an inspiring discussion of his results on Steiner symmetrization that provided the original motivation for our work, and to Bob Jerrard for pointing out an error in an earlier version of Eq. (7.5). We note that a non-convergence result similar to Proposition 6.2b was obtained independently by Bianchi, Klain, Lutwak, Yang, and Zhang [6].

REFERENCES

[1] L. V. Ahlfors, *Conformal Invariants: Topics in Geometric Function Theory*, McGraw-Hill Series in Higher Mathematics, 1973.
[2] A. Baernstein II and B. A. Taylor, *Spherical rearrangements, subharmonic functions, and *-functions in n-space*, Duke Math. J. **43** (1976), no. 2, 245–268.
[3] W. Beckner, *Sobolev inequalities, the Poisson semigroup, and analysis on the sphere $\mathbb{S}^n$*, Proc. Nat. Acad. Sci. U.S.A. **89** (1992), no. 11, 4816–4819.
[4] Y. Benyamini, *Two-point symmetrization, the isoperimetric inequality on the sphere, and some applications*, Texas Functional Analysis Seminar (1983–84), Longhorn Notes, University of Texas Press, Austin, 1984, pp. 53–76.
[5] G. Bianchi and P. Gronchi, *Steiner symmetrals and their distance from a ball*, Israel J. Math. **135** (2003), no. 1, 181–192.
[6] G. Bianchi, D. Klain, E. Lutwak, D. Yang, and G. Zhang, *A countable set of directions is sufficient for Steiner symmetrization*, Adv. in Appl. Math. **47** (2011), no. 4, 869–873.
[7] H. J. Brascamp, E. H. Lieb, and J. M. Luttinger, *A general rearrangement inequality for multiple integrals*, J. Functional Analysis **17** (1974), no. 2, 227–237.
[8] F. Brock and A. Yu. Solynin, *An approach to symmetrization via polarization*, Trans. Amer. Math. Soc. **352** (2000), no. 4, 1759–1796.
[9] A. Burchard and M. Schmuckenschläger, *Comparison theorems for exit times*, Geom. Funct. Anal. **11** (2001), no. 4, 651–692.
[10] C. Carathéodory and E. Study, *Zwei Beweise des Satzes, daß der Kreis unter allen Figuren gleichen Umfanges den größten Inhalt hat*, Math. Ann. **68** (1909), no. 1, 133–140.
[11] E. A. Carlen and M. Loss, *Extremals of functionals with competing symmetries*, J. Funct. Anal. **88** (1990), no. 2, 437–456.
[12] D. Coupier and Y. Davydov, *Random symmetrizations of convex bodies*, Preprint `arXiv:1211.1785` (2012).
[13] V. N. Dubinin, *Transformation of condensers in space*, Dokl. Akad. Nauk SSSR **296** (1987), no. 1, 18–20.
[14] R. Durrett, *Probability: Theory and Examples*, third ed., Duxbury Press, Belmont, CA, 2005.
[15] H. G. Eggleston, *Convexity*, Cambridge University Press, New York, 1958.
[16] G. B. Folland, *Real Analysis*, 2^{nd} ed., Pure and Applied Mathematics, Wiley & Sons, New York, 1999.
[17] M. Fortier, *Convergence results for rearrangements: Old and new*, M.Sc. Thesis, University of Toronto, December 2010.
[18] D. Klain, *Steiner symmetrization using a finite set of directions*, Adv. in Appl. Math. **48** (2012), no. 2, 340–353.
[19] B. Klartag, *Rate of convergence of geometric symmetrizations*, Geom. Funct. Anal. **14** (2004), no. 6, 1322–1338.
[20] B. Klartag and V. D. Milman, *Isomorphic Steiner symmetrization*, Invent. Math. **153** (2003), no. 3, 463–485.
[21] L. Lyusternik, *Die Brunn-Minkowskische Ungleichung für beliebige meßbare Mengen*, Dokl. Acad. Sci. URSS **3** (1935), 55–58.
[22] P. Mani-Levitska, *Random Steiner symmetrizations*, Studia Sci. Math. Hungar. **21** (1986), no. 3-4, 373–378.
[23] M. Meyer and A. Pajor, *On Santaló's inequality*, Geometric Aspects of Functional Analysis (1987–88), Lecture Notes in Math., vol. 1376, Springer, Berlin, 1989, pp. 261–263.
[24] C. Morpurgo, *Sharp inequalities for functional integrals and traces of conformally invariant operators*, Duke Math. J. **114** (2002), no. 3, 477–553.
[25] Y. Peres and P. Sousi, *An isoperimetric inequality for the Wiener sausage*, Geom. Funct. Anal. **22** (2012), no. 4, 1000–1014.
[26] J. Steiner, *Einfacher Beweis der isoperimetrischen Hauptsätze*, Crelle J. Reine Angew. Math. **18** (1838), 281–296.
[27] J. van Schaftingen, *Approximation of symmetrizations and symmetry of critical points*, Topol. Methods Nonlinear Anal. **28** (2006), no. 1, 61–85.
[28] J. van Schaftingen, *Explicit approximation of the symmetric rearrangement by polarizations*, Arch. Math. (Basel) **93** (2009), no. 2, 181–190.
[29] A. Volčič, *Random symmetrizations of measurable sets*, Calc. Var. PDE, to appear (2012).
[30] V. Wolontis, *Properties of conformal invariants*, Amer. J. Math. **74** (1952), no. 3, 587–606.

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF TORONTO, 40 ST. GEORGE STREET, TORONTO, ONTARIO, CANADA M5S 2E4
E-mail address: almut@math.toronto.edu, reitrof.cram@gmail.com